A Deep Learning Approach for Biometric Security in Video Surveillance System Using Gait

Naseer Rajasab*, Mohamed Rafi

Computer Science and Engineering, U.B.D.T. College of Engineering, Davanagere 577004, India

Corresponding Author Email: nasiryr@bietdvg.edu

ABSTRACT

Video surveillance systems and biometrics inclusion play a significant part in various applications like a criminal investigation, medical rehabilitation, virtual reality, etc. Human Gait is a popular biometric where the individual is differentiated by unique limb actions and special ground reaction force. The unique movement of limb actions is called gait, and a record of 2-D aggregate floor response force through one walking cycle is called Cumulative Foot Pressure Images (CFPI). Both gait and cumulative foot pressure images can be acquired simultaneously of the same person during walking under a surveillance system for human identification. Accurate gait recognition is highly impactful for most applications and a major challenge for researchers due to various external factors like different shoes, mood, clothes, injuries etc., affects the individual gait. The novel system addresses the accuracy issue and proposes two models using the Deep Convolution Neural Network (DCNN) architecture on a large standard database, CASIA-D, containing gait pose and CFPI images of the same person. First, the model of the DCNN is trained using unique Gait Energy Image (GEI) features which reduce the computational time compared to other types of feature sets. The second model is prepared using the CFPI features. Experimentation has been carried out to evaluate the performance of these models with different optimization methods and activation functions and has proven that the DCNN model is far superior in building an accurate gait recognition system on large standard datasets.

1. INTRODUCTION

Human biometric recognition schemes are used in a variety of applications, including banking, airports, the military, and public safety. In a variety of industries, including e-commerce, internet access, physical access control, PDA, government applications like the national ID card, social security, welfare disbursement, border control, and military surveillance, biometrics has grown and established itself as a trustworthy source. The use of biometric systems is expanding in both the private and public sectors, which is creating significant security holes in the systems that have already been established. Currently, using biometric technology for human recognition is the most efficient method. Biometrics is utilised for recognising and validating individuals based on physiological and behavioural characteristics. Biometrics is separated into two portions via physiological and behavioural features. Physiological traits are related to the organisation of the body. These contain limited instances such as fingerprints, Iris scans, footprints, etc. Behavioral features are linked to the behavioural designs of a person. These include limited models such as gait, speech designs, signatures, etc. Here, we utilised behavioural features such as gait for human biometric recognition. Gait detection is an alternative method for authenticating a person based on how they walk. Gait detection has numerous benefits compared to other biometric procedures. First, it works well when people are at a distance. Second, it requires low-resolution images as compared to face recognition. Third, it does not require any co-operation from the subject compared to other biometric recognition systems.

Fourth, it works well when other biometric features are hidden, like the face and iris. Lastly, it is very difficult to imitate the gait of a person.

Human beings walk through the actions of their lower limbs, like striding or sprinting. The order of these actions comprises the individual gait. One walking pattern is constructed from the posture change and oscillation stages executed by the legs. Gait shape is determined by biomechanical and energetic features [1].

Additionally, every person possesses an exclusive gait signature utilised for recognition functions in a gait recognition procedure [2]. An exciting feature of gait detection is that it can recognise individuals without their awareness or consent, conflicting with more biometric approaches like face detection. Visual-founded techniques control the area of walking-style detection [3, 4]. Investigators can take advantage of many effortlessly accessible video databases [5], which comprise over 10,000 people [6]. Current developments display an identification degree of between 90% and 95% in ideal observing circumstances; nonetheless, precision reduces in difficult cases (like obstructions, view differences, or look variations) [4]. Wearable inertial sensors are projected to identify gait [7, 8]. These detectors are used a lot in bioscience to study how people walk, and as a result, they help a lot with research [9]. Cumulative foot pressure image comprises increasing spatial and temporal force info through one gait cycle [10, 11], which might aid in managing the problems in identifying the diverse shoes-wearing person. For a few safety situations like jailhouse safety structure, bathhouse, entry at community transport, and entry at Japanese house, camera-

based detection structure doesn't function fine.

The gait recognition system is affected by various environmental factors such as walking speed, clothing, mood, multiview, and different shoes Due to this, accurate gait recognition is a challenge and is important for many researchers. To identify people and let them into a restricted area, biometric-based techniques have emerged as a promising new option [12]. Gait recognition is used in many areas, like home security and public places, to figure out what people are doing. It is also used in medicine to figure out what kind of gait disease someone has, like Parkinson's or neuropathic. In recent years, it has also been used in the animation and gaming industries to make virtual versions of natural walking styles.

Advanced machine learning and pattern recognition techniques may now be implemented using the latest deep learning techniques, such as Deep Convolutional Neural Networks (DCNNs). Using CNN, human gait recognition has never been studied previously. It has been used successfully to handle a variety of difficult recognition issues. We describe a recognition strategy for inferring complex non-linear capabilities using high-dimensional pictures and a CNN architecture. Convolution followed by subsampling is typical in convolutional neural networks. The element of the convolutional layer uses a specific channel to delineate the picture (weights). Many of these layers are abruptly merged into a deep convolutional organization. A completely connected neural system layer performs the final characterization in the last stage. The final stage feeds it. Note that CNN can extract visual samples from photos with little preprocessing. Deep CNNs for step recognition will be the subject of this research, which aims to offer essential engineering. We perform several experiments to determine how many convolutional, subsampling, and fully related layers are best. Observation is also used to establish the appropriate number of component maps per layer. Lastly, we use DCNN on the CASIA-D database to test the quality and feasibility of our proposed work, and we compare the results to existing state-of-the-art approaches in step acknowledgment analysis [13].

2. LITERATURE REVIEW

Davarzani et al. [14] presented a deep learning-type system for detecting a person walking. They started with three models: linear regression, artificial neural network (ANN), and LSTM. The results of these systems were then pooled for the results. Using publicly available datasets, the training was carried out with each design. The findings showed that the ANN performed well compared to linear regression and LSTM. In the experiment, three alternative viewpoints of the CASIA B dataset were used.

Anusha and Jaidhar [15] exhibited a modified Local Optimal Oriented Pattern Binary (MLOOPB). MLOOPB descriptors were a development of LOOP descriptors. The variety of features recovered from MLOOPB, including the histogram and horizontal width variables, The retrieved features were then decreased using a novel technique before categorization. The experiment was carried out on a standard dataset, like CASIA-B, and the results prove that the experimental approach was more accurate than the other techniques.

Connor [16] presented foot force information gathered from 92 participants moving on a pressure mat, either unshod or shod. The aim was to find the identification and detection performance of the recognising system in three cases: barefoot, same shoe, and different shoes. Connor then evaluated a variety of pressure-derived variables and general walking-style parameters. The output gives the classification accuracy of 86.5 percent, which was achieved for the most difficult case.

Wan et al. [17] surveyed study functions in gait detection. Additionally, detection is founded on video; novel modes like detection based on on-ground sensors, radars, and accelerometers; novel methods which comprise machine learning systems; and inspection tasks and susceptibilities in this area. They also discovered a list of upcoming investigation instructions. The evaluation discloses present hi-tech and obliges both specialists and novices in gait detection. Furthermore, this catalogues upcoming compositions and openly accessible files in gait detection for investigators.

Rafi et al. [18] presented a model-founded method for gait detection utilising the scientific model of geometry and image dispensation methods. In such a method, characteristic matrices used for gait detection are built utilising division, Hough convert, and corner recognition methods. Certainly, this is likely to identify a person by examining the gait restrictions extracted by their footsteps in diverse settings. In the preprocessing phase, picture frames captured by video systems are entered into the Canny Edge recognition algorithm to sense image boundaries and decrease clatter from Gaussian filtering. Later, the Hough transform is applied to separate characteristics of preprocessing outcomes and to acquire gait templates. The Gait Stricture Final is utilised for extracting gait strictures, and the Harris Corner Recognition method is used to detect edges and generate characteristic points. Gait parameters are calculated using distinct points and later kept in the gait database. Using the gait detection interface, arbitrary subjects' strictures compete alongside model group inaccessible databases for detection. The projected technique has deemed a database counting ten topics and five constraints as a gait recognition scheme. It is important to note that when the camera is positioned at 90 and 270 degrees toward the individual, each detection constraint is noticeable, quantifiable, and requires more than 80% precision in detection outcomes.

Wang et al. [19] presented an easy but effective gait detection algorithm utilising spatial-temporal outline examination. For every image order, background subtraction algorithm and easy correspondence process are primarily used for section and follow moving outlines of striding person. Later, eigenspace transformation founded on Principal Component Analysis (PCA) is used to time-changing space signs resulting by order of outline images for reducing the dimensionality of input characteristic space. Experimental pattern cataloguing methods are lastly achieved in lower-dimensional eigenspace for detection. This technique indirectly seizes structural and transitional features of gait. Widespread practical outcomes on outdoor image orders exhibit the projected algorithm's inspiring detection performance with a comparatively short computational price.

Wolf et al. [20] presented a 3-Dimensional CNN with a 3-Dimensional spatiotemporal tensor as input, containing the grayscale image for the primary network and optical movement for 2nd and 3rd networks. The prototype skills plus tried utilising the CASIA-B dataset, MoBo database, and UFS database. The method was assessed on differences in striding rapidity, attrition, and view angle.

Following the detailed literature review, the following research challenges have been identified.

- To test the performance of the algorithms, the majority of

the researchers used publicly available standard datasets like USF, CMU, CASIA-A, B, and C. There has been little study done on the CASIA-D dataset (gait-footprint dataset).

- A few external aspects, such as views and carrying bag conditions, have been handled; nevertheless, more exploration and investigation into the larger dataset is still open.
- As many indicators as possible should be combined or fused in meaningful ways to improve gait recognition performance. Different gait traits or biomechanics can be used as hints. The fusion-based technique can boost detection capability. Thus, there is a need for more research.

2.1 Literature review summary

It is evident from the summary that majority of the research were conducted on smaller private datasets. Therefore, the efficiency of the suggested method is conducted on publicly available benchmark dataset CASIA-D. Besides, there are several drawbacks of the gait recognition methodologies of the aforementioned related works. First, dynamic behavioral features were extracted considering a specific set of body joints which might not be sufficiently differentiating. Second, gait recognition methodologies were limited to conventional machine learning models, which depends handcrafted features. This is undesirable due to the involvement of heavy computation. The extraction of handcrafted features requires careful selection and domain-specific knowledge, which might be difficult to attain. In this article, all the above limitations are addressed to overcome the dependency on a specific set feature.

The gait recognition method is further improved by avoiding domain-specific handcrafted feature extraction to mitigate the difficult process of feature engineering and feature selection using unique DCNN architecture. The use of DCNN architecture allows extracting optimized hierarchical discriminating features without the necessity of the manually extracted features. Thus, the designed architecture can be applied not only to gait but also to other biometric datasets. This enhances the portability and limits the dependency on a specific dataset. It also focuses on analysis of previous research on gait recognition. The challenges are identified. The deep learning approaches and their use in biometric domain are provided. This provides motivation to develop new method of the gait recognition using deep learning. To fill the research gaps of the prior researches and to enhance the accuracy of the gait recognition, deep learning-based approaches are proposed in the subsequent sections.

3. PROPOSED METHODOLOGY

The proposed research design is provided in Figure 1. It consists of two phases: training and testing. Both the phases consist of the same number of steps. In the first step, the gait-footprint images are acquired from the standard repository. Based on the training and testing ratio of 80:20, the images are separated into labeled(training) and unlabeled images(testing). After that, unwanted noise has been removed and images have been converted into model-free representations like Gait Energy images for feature extraction. In a later step, the model is trained using the DCNN architecture (VGG-19). The dataset contains 88 class labels as subject_ID. Finally, the images have been classified using a class label for the recognition of humans in surveillance systems.

Gait biometrics are extensively researched for verification and access control [21]. In the current scenario, deep learning has attained huge achievements in the areas of safe computing [22, 23], and activity detection [24]. Deep learning-based gait identification approaches also increased functional connectivity over conventional machine learning-based techniques, e.g., SVMs [25]. Since the excellent capability of DCNNs in image-feature abstraction, numerous scientists have used DCNNs for gait or action detection [26].

3.1 Dataset

The dataset is obtained from the standard CASIA repository [27]. The CASIA-D dataset was composed of twenty female participants and sixty-six male individuals. All subjects are Asian men and women between the ages of 20 and 60. The video frames or gait pose images are captured through the camera, and the cumulative foot pressure images are captured through the flooring mat sensors, as shown in Figure 2. A copy of the dataset is shown in Figure 2, along with its distribution over age and body mass index (BMI). The participants are told to walk normally (5 times) and quickly (5 times) through the pressure sensor. There are 10x88 = 880 cumulative foot pressure recordings, each with three cumulative foot pressure images, for a total of 2640 images to analyze when examining the effect of walking speed. Similarly, for every four-gait pose, there are 3520 images; for every record size, it becomes 6160(3520+2640) images. There are two file formats, one for the gait-pose images and another for the CFP images. These datasets include both gait and foot pressure trials. The file format for each foot pressure image is as follows: (left(0), right(1), first footwear (2), second footwear (3))_(data index, i). Gait-pose images have a filename format [subject-Id]. The metadata for each trial has the following format: [subject ID, mass, age, shoe size, gender]. The images are stored in.png format with a resolution of 173x353.

3.2 Pre-processing

After the data has been collected, the images must be processed to make the data they show easier to understand while keeping important information that can be used in the feature extraction and classification steps and making new images that can be used as input in those steps. Filtering is a fundamental function that can accomplish tasks like noise reduction. The Gaussian smoothing filter is regarded as the "ideal" blur for many applications. For the noise suppression produced by the illumination, we employed the Gaussian Filter. The one-dimensional Gaussian distribution is transformed into a two-dimensional filter.

$$G(x) = \frac{1}{\sqrt{2\pi\alpha}} e^{-\frac{x^2}{2\alpha^2}} \tag{1}$$

where, σ is the distribution's standard deviation.

3.3 Background subtraction

Background removal is a typical method for detecting moving objects in video surveillance systems. Segmenting moving objects leverages the difference between the backdrop and input pictures. In the first step, called "backdrop initialization," the background image is taken from a certain point in the video sequence. In the second step, the background is updated due to changes in the real scene and the frame difference.
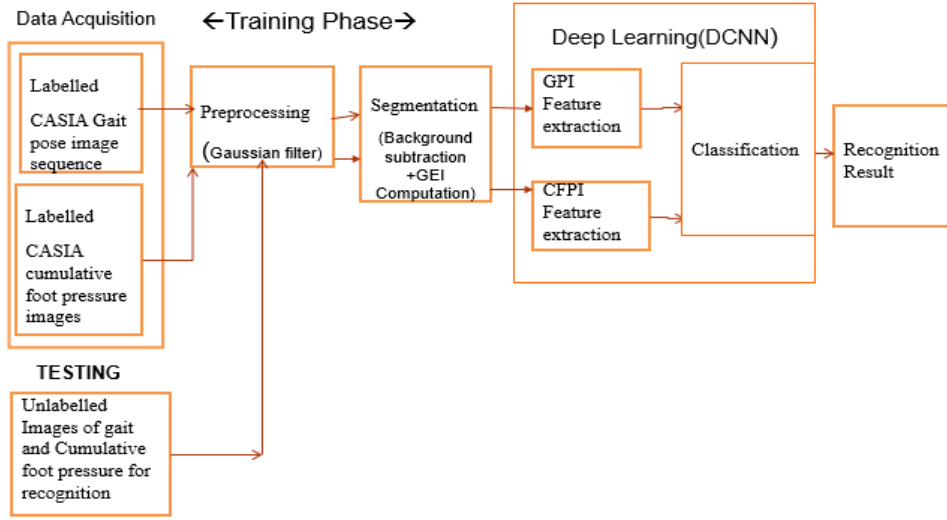
Figure 1. Proposed system for human identification in video surveillance system
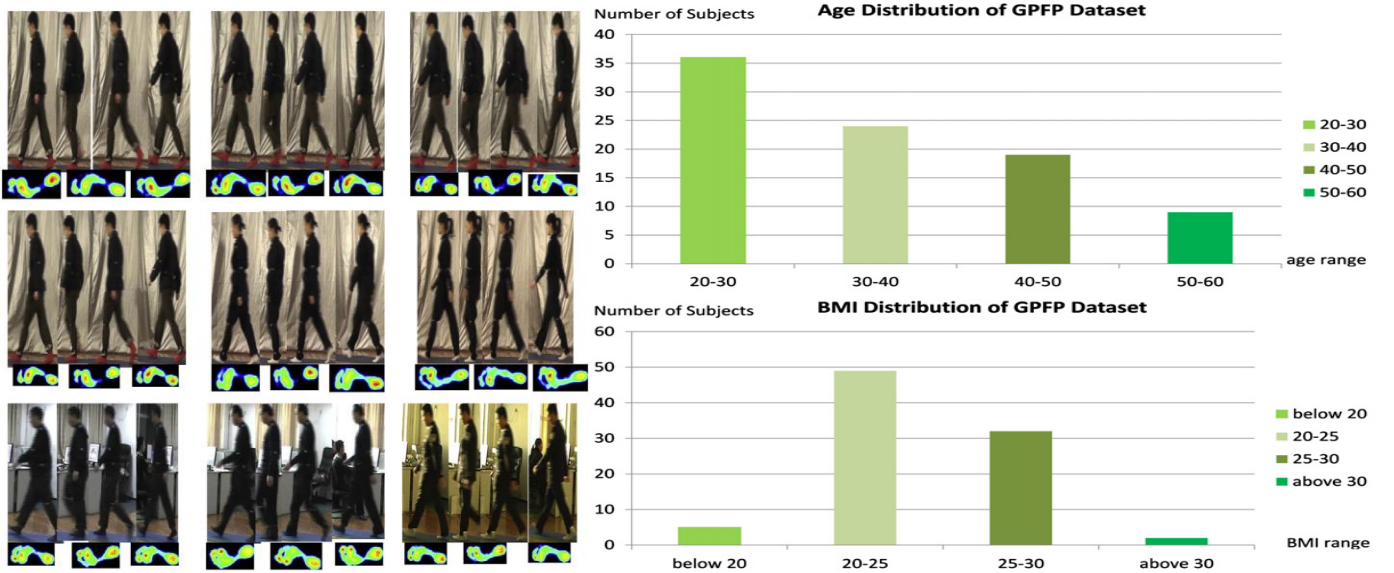
Figure 2. Sample dataset and chart representing distribution of CASIA-D (Gait-footprint) dataset over age and BMI (Body Mass Index) [11]

Background subtraction and frame difference were used in this investigation. To tell a moving object from its background, take out the background and leave only the moving object or foreground.

As a result, in this study, the moving item or human detection might be characterized as follows:

$$H(x) = \sum_{i=1}^{n} f(x,j) \qquad (2)$$

$$H(y) = \sum_{i=1}^{m} f(i,y) \qquad (3)$$

Thresholding on H(x) and H(y) is used to conduct background subtraction or frame difference (y).

3.4 Gait energy image computation

GEI is the average image of all binary silhouette frames captured during a gait cycle. The following formula is used to construct a GEI G from a series of size normalized and aligned 2D silhouettes B:

$$G(x,y) = \frac{1}{N} \sum_{i=1}^{N} B_t(x,y) \qquad (4)$$

where, x and y are two-dimensional coordinates, N is the total number of frames in the sequence, and t is the frame index After the GEI Computation, the following Gait Energy Image is obtained as shown in Figure 3.

Figure 3. Different binary silhouettes and corresponding gait energy image
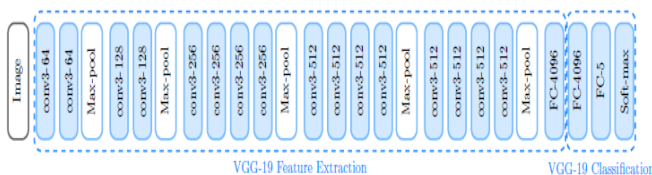
3.5 Feature extraction

The Multilayer Perceptron Network (DCNN) is designed for processing, averaging, and standardising layers with two consecutive triples and fully connected layers with two subsequent layers fed a GEI image for training the people independently.
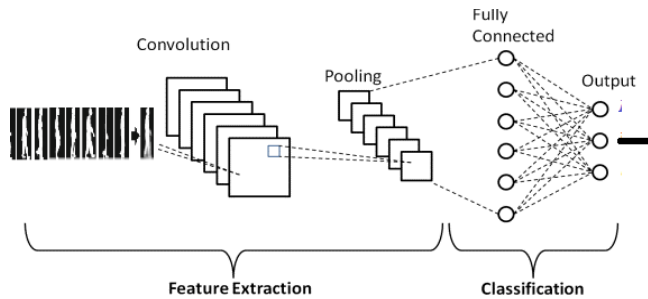
Figure 4. Deep convolutional neural network

Following image preprocessing and acquiring the appropriate input images, the feature extraction module produces a feature map for walking pose photos, as illustrated in Figure 4, which will subsequently be used for classification. The VGGNet-19 was chosen as the architecture for this task as shown in Figure 5, and it was implemented in Python using the Keras deep learning library. VGGNet is a multilevel, deep convolutional neural network (CNN) structure for the Visual Geometry Group. VGGNet-16 or VGGNet-19 has 16 or 19 convolutional layers, respectively, and the term "depth" refers to the number of layers. It is a very deep CNN that has recently done well in many tasks involving recognizing and sorting images.

Other networks, such as ResNet and GoogleNet, performed better in this classification competition with. On the other hand, VGG networks are easier to implement and train since they converge faster, and they are generally the best solution for feature extraction.

Figure 5. Architecture of VGG-19 network

3.6 Classification

After extracting the Gait Pose Images features using DCNN(VGG-19), the model has been trained on the VGG-19 classifier, and model performance metrics have been recorded for the evaluation, as shown in Figure 6.
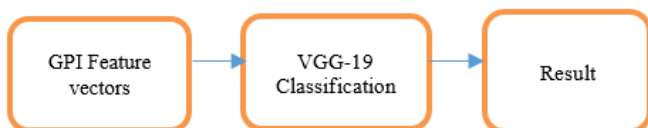
Figure 6. Video-based gait or GPI classification step

The deep learning architectures discussed in the previous sections are sometimes utilized for extracting features from the input dataset rather than for classification. In this case, the feature vectors are then put into groups using more traditional machine learning methods.

After extracting the cumulative foot pressure image features using DCNN(VGG-19), the model was trained on a VGG-19 classifier. Model performance metrics were recorded for the evaluation to compare with the DCNN in terms of accuracy, as shown in Figure 7.

Figure 7. CFP images classification step

4. RESULT AND DISCUSSION

Experimental results of video-based gait and cumulative foot pressure for human identification in the video surveillance system are given below:

4.1 Video-based gait recognition system

The following metrics have been calculated to see the Model's performances. It is recorded and mentioned in Table 1 below.

Table 1. Performance metrics of video-based gait model

Subject ID	Precision	Recall	F1-Score
1	0.9656	0.9527	0.9602
2	0.9649	0.9649	0.9584
3	0.9536	0.9543	0.9622
4	0.9554	0.9644	0.9556
5	0.9618	0.9554	0.9450
6	0.9687	0.9611	0.9555
7	0.9562	0.9697	0.9552
8	0.9507	0.9522	0.9753
9	0.9622	0.9622	0.9526
10	0.9543	0.9758	0.9564
Accuracy			96.28
Macro avg	0.9634	0.9622	0.9625
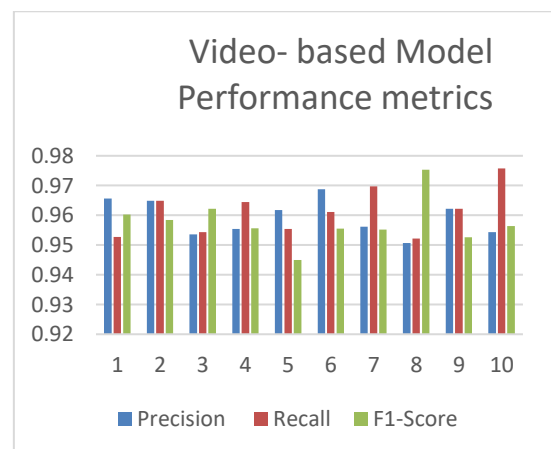
Figure 8. Plot representation of model performance

The detection performance of each test set is determined, and the average detection rate is displayed. This study reports the specificity, recall, and F-score assessments by calculating the macro-mean for each class. To get the macro-mean, the specificity, recall, and F-score metrics for each category are calculated, and the associated weighted average is found.

The equivalent graph representation is shown in Figure 8.

$$Accuracy = \frac{True\ Positive + True\ s}{Total\ Sample} \quad (5)$$

$$Precison = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (6)$$

$$F1 = 2 * \frac{1}{\frac{1}{precision} + \frac{1}{recall}} \quad (7)$$

4.2 Learning curves of training and validation

The loss in training and validation over epochs shows if the model is generalising or memorizing. The model learns a generalised pattern if the training and validation losses gradually reduce across epochs. The learning curve can thus be used to identify model overfitting. Figure 10 illustrates the suggested CNN model's average learning and verification loss on the CASIA D dataset due to a cross-validation experiment. Learning and verification loss diminish as learning and verification accuracy eventually improve. Figure 9 depicts the training and validation accuracy of the CASIA-D dataset over time.

a) Accuracy curve   b) Loss curve
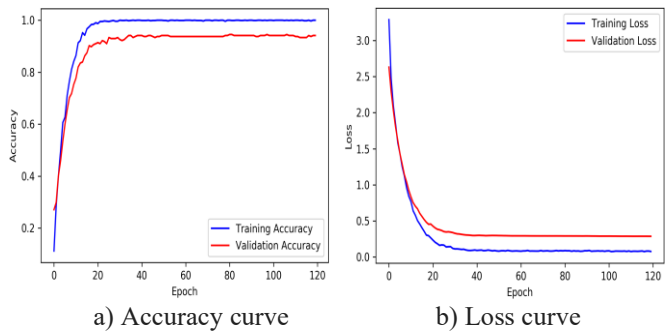
Figure 9. Average learning curve of the proposed work on CASIA video-based gait dataset

4.3 Cumulative foot pressure images based gait recognition

The model's performance has been calculated, recorded, and mentioned in the Table 2. The detection performance of each test set is determined, and the average detection rate is displayed. This study reports the specificity, recall, and F-score assessments by calculating the macro-mean for each class. To get the macro-mean, the specificity, recall, and F-score metrics for each category are calculated, and the associated weighted average is found. The equal graph representation is shown in Figure 10.

4.4 Learning curves of training and validation

On the CFPI (CASIA-D) dataset, Figure 10 displays the proposed model's performance. It can be shown in Figure 11 that the direction of validation loss over epochs is downward. Furthermore, once validation loss reaches a plateau, it does not

grow, and there is no overfitting. On the CFPI dataset, the difference in training and validation accuracy is small.

Table 2. Performance metrics of cumulative foot pressure based gait model

S.Id	Precision	Recall	F1-Score
1	0.9157	0.9026	0.9103
2	0.8748	0.8844	0.8785
3	0.8837	0.8642	0.8826
4	0.8655	0.8848	0.8857
5	0.8819	0.8758	0.8753
6	0.8888	0.8818	0.8854
7	0.8763	0.8798	0.8753
8	0.8604	0.8725	0.8754
9	0.8828	0.8825	0.8726
10	0.8747	0.8555	0.8867
			88.31
Macro avg	0.8860	0.8845	0.8755
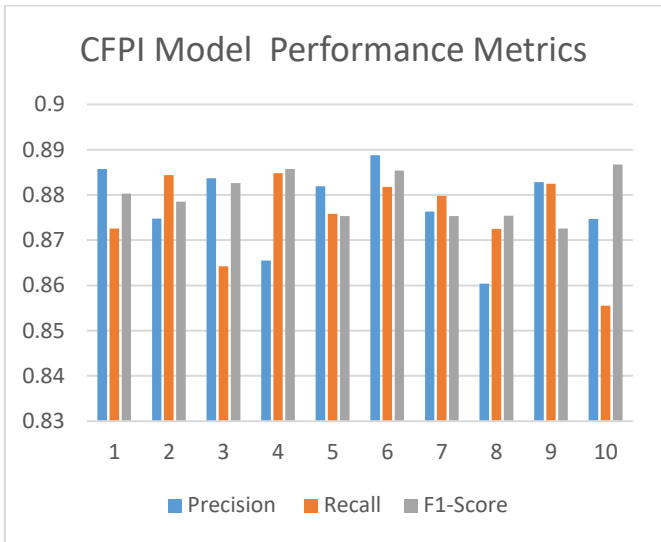
Figure 10. Plot representation of model performance


a) Accuracy curve   b) Loss curve
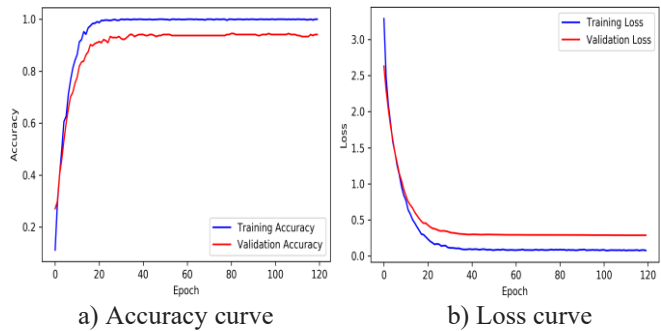
Figure 11. Average learning curve of the proposed work on CASIA foot pressure image dataset

4.5 Hyper-parameter tuning of CNN

The VGG-19 CNN network was chosen for feature extraction and classification (Figure 5). Only the fully connected layers were trained during the hyper-parameter tuning step. While the weights of the convolutional levels remained constant, the effects of relearning extra layers will be investigated next. Based on the RAM available on the graphic card, the batch size was set to 16. A greater value would help the model converge faster because more data

would be used to update the network in each training epoch, but it might cause problems when allocating too much memory. The number of epochs was chosen to be 35 since it was discovered that training accuracy quickly converged towards 100 percent and that further training would result in overfitting. Also, an early stopping approach was used, in which the training executes for the specified epochs. Still, the registered system is the first to get the highest validation accuracy during training.

The whole dataset was used for training and validation, using 10-fold cross-validation and the folds separated by subjects. The learning rate was changed to flatten out the accuracy curve even further. Finally, 0.0002 was chosen as the 10-fold cross-validation learning rate value. The number of epochs was also increased to 50 to ensure that no under-fitting occurred because the learning rate was reduced, resulting in a slower convergence time as shown in Figure 12. It's worth mentioning that the validation accuracy and loss stabilise early in training. Also, the training accuracy is close to 100%, which shows that the model is overfitted and that this chain of epochs is right.
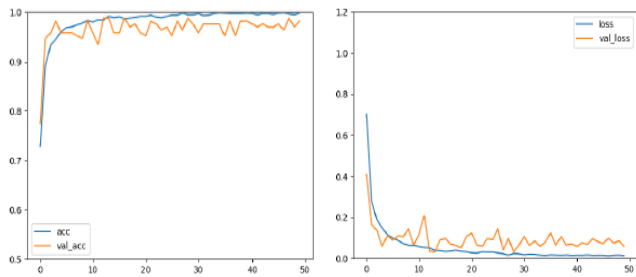
Figure 12. Learning (blue) and verification (orange) correctness (left) and losses (right)

4.6 Performance of activation function and enhancement method

In a series of experiments, the suggested model is compared to various activation functions and optimization methodologies as shown in Table 3. The max-out actuates functions are suitable for deep training networks, and a powerful system can be created using a max-out network with a dropout regularizer. It was chosen for comparison. Using different optimization methods, the objective function is minimized.

Table 3. On the CASIA dataset, mean detection accuracy of the proposed deep system with various activation functions and enhancement strategies

Suggested features + neural network	Accuracy	Specificity	Recall	F-score
Maxout network	95.28	96.32	94.44	94.40
Suggested DLNN + ReLU + RMSProp	95.32	95.55	95.22	95.40
Suggested DLNN + tanh + RMSProp	95.44	95.66	94.17	94.40
Suggested DLNN + ReLU + Adam	96.11	95.25	95.22	96.40
Suggested DLNN + tanh + Adam	96.28	97.10	96.62	96.55

4.7 Comparison of DCNN accuracy with other deep learning models

The proposed DCNN's recognition accuracy is compared with different methods as shown in Table 4 to reveal that the suggested residual learning-based CNN architecture extracts more distinctive features than handcrafted features. The network training is more effective than in previous studies. The equivalent graph representation is shown in Figure 13.

Table 4. Comparison analysis of different deep learning algorithms with proposed work for video-based gait images

Sl. No.	Reference	Year	Method	Accuracy
1	Proposed Work	2022	DCNN	96.28
2	Chao et al. [28]	2020	Partial RNN	86.5
3	Sepas-Moghaddam and Etemad [29]	2020	GaitPart	88.5
4	Hou et al. [30]	2020	GLN	89.5
5	Li et al. [31]	2020	HMRGait	89.5
6	Lin et al. [32]	2020	3D CNNGait	90.4
7	Yao et al. [33]	2019	PoseGait	74.9
8	Sokolova and Konushin [34]	2019	DisentangledGait	79.9
9	Nair and Kendricks [35]	2017	DBNGait	60.7
10	Wu et al. [36]	2017	CNNGait	73.5
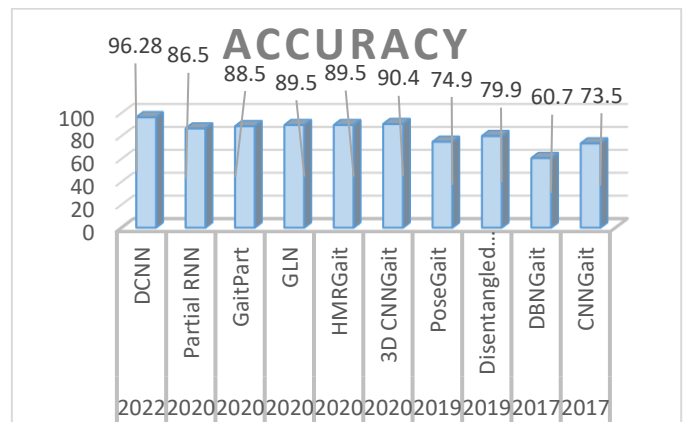
Figure 13. Accuracy comparison graph

5. CONCLUSION AND FUTURE WORK

In this paper, a deep learning strategy for accurate gait recognition for the CASIA-D database is presented. This larger standard database contains a variety of considerations considering the various factors like different shoes, normal walking, fast walking etc. On huge datasets, the accuracy obtained is the best among the classical classifiers studied in this work. The recognition accuracy of the proposed methods on huge datasets is higher than in all prior researches. Numerous examples of successful integration of the latest biometric technologies in practice can be found in the domains of information security, surveillance, medicine, finance, education, retail, and others. One of the first domains that was fundamentally transformed by the introduction of deep learning architectures was computer vision. As we have used gait energy image feature representation technique to train the

DCNN structure it may reduce the performance due to change in the viewing angle. In the future, the fusion method can be used by combining the above two models to build a robust gait recognition system that can help adapt to the real-world scenario.

REFERENCES

[1] Holt, K., Jeng, S.F., Ratcliffe, R., Hamill J. (1995). Energetic cost and stability during human walking at the preferred stride frequency. J. Mot. Behav., 27(2): 164-178. https://doi.org/10.1080/00222895.1995.9941708

[2] Connor, P., Ross, A. (2018). Biometric recognition by gait: A survey of modalities and features. Computer Vision and Image Understanding, 167: 1-27. https://doi.org/10.1016/j.cviu.2018.01.007

[3] Rida, I., Almaadeed, N., Almaadeed, S. (2018). Robust gait recognition: A comprehensive survey. IET Biom., 8(1): 14-28. https://doi.org/10.1049/iet-bmt.2018.5063

[4] Singh, J.P., Jain, S., Arora, S., Singh, U.P. (2018). Vision-based gait recognition: A survey. IEEE Access, 6: 70497-70527. https://doi.org/10.1109/ACCESS.2018.2879896

[5] Makihara, Y., Matovski, D.S., Nixon, M.S., Carter, J.N., Yagi, Y. (2015). Gait Recognition: Databases, Representations, and Applications. In Wiley Encyclopedia of Electrical and Electronics Engineering; John Wiley & Sons, Inc.: Hoboken, NJ, USA, pp. 1-15. https://doi.org/10.1002/047134608X.W8261

[6] Takemura, N., Makihara, Y., Muramatsu, D., Echigo, T., Yagi, Y. (2018). Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. IPSJ Transactions on Computer Vision and Application, 10: 4. https://doi.org/10.1186/s41074-018-0039-6

[7] Gafurov, D., Snekkenes, E. (2009). Gait recognition using wearable motion recording sensors. EURASIP Journal on Advances in Signal Processing, 2009: 415817. https://doi.org/10.1155/2009/415817

[8] Springer, S., Juric, M.B. (2015). Inertial sensor-based gait recognition: A review. Sensors, 15(9): 22089-22127. https://doi.org/10.3390/s150922089

[9] Vienne, A., Barrois, R.P., Bu_at, S., Ricard, D., Vidal, P.P. (2017). Inertial sensors to assess gait quality in patients with neurological disorders: A systematic review of technical and analytical challenges. Front Psychol., 8: 817. https://doi.org/10.3389/fpsyg.2017.00817

[10] Moustakidis, S.P., Theocharis, J.B., Giakas, G. (2008). Subject recognition based on ground reaction force measurements of gait signals. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38(6): 1476-1485. https://doi.org/10.1109/TSMCB.2008.927722

[11] Zheng, S., Huang, K. (2012). A cascade two-modality fusion scheme for pedestrian identification. Pattern Recognition, 45: 3603-3610.

[12] Mouafo, D., Biaou, U. (2022). Face recognition system for control access to restrictive domain. International Journal of Safety and Security Engineering, 12(2): 251-257. https://doi.org/10.18280/ijsse.120214

[13] Nithyakani, P., Shanthini, A., Ponsam, G. (2019). Human gait recognition using deep convolutional neural network. 2019 3rd International Conference on Computing and Communications Technologies (ICCCT). https://doi.org/10.1109/ICCCT2.2019.8824836

[14] Davarzani, S., Saucier, D., Peranich, P., Carroll, W., Turner, A., Parker, E., Middleton, C., Nguyen, P., Robertson, P., Smith, B. (2020). Closing the wearable gap—Part VI: "Human gait recognition using deep learning methodologies". Electronics, 9: 796.

[15] Anusha, R., Jaidhar, C. (2020). Clothing invariant human gait recognition using modified local optimal pattern binary descriptor. Multimedia Tools and Applications, 79: 2873-2896. https://doi.org/10.1007/s11042-019-08400-8

[16] Connor, P.C. (2015). Comparing and combining underfoot pressure features for shod and unshod gait biometrics. In Proceedings of the 2015 IEEE International Symposium on Technologies for Homeland Security (HST), Waltham, MA, USA, pp. 1-7. https://doi.org/10.1109/THS.2015.7225338

[17] Wan, C.S., Wang, L., Phoha, V.V. (2018). A survey on gait recognition. ACM Computing Surveys, 51(5): 1-35. https://doi.org/10.1145/3230633

[18] Rafi, M., Khammari, H., Wahidabanu, R., Taj, Y. (2013). Model-based approach for gait recognition system. International Journal of Soft Computing and Engineering (IJSCE), 3(5): 223.

[19] Wang, J.J., Yang, J.C., Yu, K., Lv, F.J., Huang, T., Gong, Y.H. (2010). Locality-constrained linear coding for image classification. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2010.5540018

[20] Wolf, T., Babaee, M., Rigoll, G. (2016). Multi-view gait recognition using 3D convolutional neural networks. Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 4165-4169. https://doi.org/10.1109/ICIP.2016.7533144

[21] Sprenger, S., Juric, M.B. (2015). An efficient HOS-based gait authentication of accelerometer data. IEEE Transactions on Information Forensics and Security, 10(7): 1486-1498. https://doi.org/10.1109/TIFS.2015.2415753

[22] Ren, K., Wang, Q., Wang, C., Qin, Z., Lin, X. (2019). The security of autonomous driving: Threats, defenses, and future directions. Proceedings of the IEEE, 108(2): 357-372. https://doi.org/10.1109/JPROC.2019.2948775

[23] Zhao, L., Wang, Q., Zou, Q., Zhang, Y., Chen, Y. (2020). Privacy-preserving collaborative deep learning with unreliable participants. IEEE Transactions on Information Forensics and Security, 15(1): 1486-1500. https://doi.org/10.1109/TIFS.2019.2939713

[24] Ji, S., Xu, W., Yang, M., Yu, K. (2013). 3D convolutional neural networks for human action recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1): 221-231. https://doi.org/10.1109/TPAMI.2012.59

[25] Man, J., Bhanu, B. (2006). Individual recognition using gait energy image. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(2): 316-322. https://doi.org/10.1109/TPAMI.2006.38

[26] Wu, Z., Huang, Y., Wang, L., Wang, X., Tan, T. (2017). A comprehensive study on cross-view gait-based human identification with deep scans. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(2): 209-226. https://doi.org/10.1109/TPAMI.2016.2545669

[27] CASIA NLPR. Casia foot pressure image database. http://www.cbsr.ia.ac.cn/english/, accessed on 11 Nov.

2021.

[28] Chao, H., He, Y., Zhang, J., Feng, J. (2019). Gaitset: Regarding gait as a set for cross-view gait recognition. AAAI Conference on Artificial Intelligence, Honolulu, HW, USA.

[29] Sepas-Moghaddam, A., Etemad, A. (2021). View-invariant gait recognition with attentive recurrent learning of partial representations. IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(1): 124-137. https://doi.org/10.1109/TBIOM.2020.3031470

[30] Hou, S., Cao, C., Liu, X., Huang, Y. (2020). Gait lateral network: Learning discriminative and compact representations for gait recognition. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science(), vol 12354. Springer, Cham. https://doi.org/10.1007/978-3-030-58545-7_22

[31] Li, X., Makihara, Y., Xu, C., Yagi, Y., Yu, S., Ren, M. (2020). End-to-end model-based gait recognition. In: Ishikawa, H., Liu, CL., Pajdla, T., Shi, J. (eds) Computer Vision – ACCV 2020. ACCV 2020. Lecture Notes in Computer Science(), vol 12624. Springer, Cham. https://doi.org/10.1007/978-3-030-69535-4_1

[32] Lin, B., Zhang, S., Bao, F. (2020). Gait recognition with multiple-temporal- scale 3D convolutional neural network. ACM International Conference on Multimedia, Seattle, WA, USA, pp. 3054-3062. https://doi.org/10.1145/3394171.3413861

[33] Yao, L., Kusakunniran, W., Wu, Q., Zhang, J., Tang, Z. (2018). Robust CNN-based gait verification and identification using skeleton gait energy image. 2018 Digital Image Computing: Techniques and Applications (DICTA), Canberra, Australia. https://doi.org/10.1109/DICTA.2018.8615802

[34] Sokolova, A., Konushin, A. (2019). Pose-based deep gait recognition. IET Biometrics, 8(2): 134-143. https://doi.org/10.1049/iet-bmt.2018.5046

[35] Nair, B.M., Kendricks, K.D. (2016). Deep network for analyzing gait patterns in the low-resolution video towards threat identification. Electronic Imaging, 2016(11): 1-8. https://doi.org/10.2352/ISSN.2470-1173.2016.11.IMAWM-471

[36] Wu, Z., Huang, Y., Wang, L., Wang, X., Tan, T. (2017). A comprehensive study on cross-view gait-based human identification with deep CNNs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(2): 209-226. https://doi.org/10.1109/TPAMI.2016.2545669

NOMENCLATURE

GEI	Gait Energy Image
CFPI	Cumulative Foot pressure Images
DCNN	Deep Convolutional Neural Network
GPI	Gait Pose Images
CASIA	Chinese Academy of Sciences' Institute of Automation